PRIMARY AND SECONDARY SUPERRESOLUTION BY DATA INVERSION (PREPRINT)

Charles Matson
David W. Tyler

6 June 2005

Journal Article

AIR FORCE RESEARCH LABORATORY
Directed Energy Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)* 6-Jun-05	2. REPORT TYPE Journal Article		3. DATES COVERED *(From – To)* Dec-5-01 to Sep-30-06
4. TITLE AND SUBTITLE Primary and Secondary Superresolution by Data Inversion (Preprint)			**5a. CONTRACT NUMBER** F29601-01-D-0083 TO 06
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER 69120G
6. AUTHOR(S) Charles L. Matson, David W. Tyler			**5d. PROJECT NUMBER** 2304
			5e. TASK NUMBER B3
			5f. WORK UNIT NUMBER AA
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) ARFL/DESA 3550 Aberdeen Ave. SE Kirtland AFB, NM 87117-5776			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) AIR FORCE RESEARCH LABORATORY 3550 Aberdeen Ave. SE Kirtland AFB, NM 87117-5776			**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/DESA
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-DE-PS-JA-2007-1006

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution is Unlimited.

13. SUPPLEMENTARY NOTES

Submitted for publication in Optics Express 14, pp. 456-473 (2005), by Optical Society of America.   GOVERNMENT PURPOSE RIGHTS

14. ABSTRACT

Superresolution by data inversion is the extrapolation of measured Fourier data to regions outside the measurement bandwidth using postprocessing techniques. Here we characterize superresolution by data inversion for objects with finite support using the twin concepts of primary and secondary supperresolution, where primary superresolution is the essentially unbiased portion of the superresolution data and secondary superresolution is the remainder. We show that this partition of superresolution into primary and secondary components can be used to explain why some researchers believe that meaningful superresloution is achievable with realistic signal-to-nose ratios, and other researchers do not.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Charles Matson
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	SAR	20	**19b. TELEPHONE NUMBER** *(include area code)*

Primary and secondary superresolution by data inversion

Charles L. Matson
Starfire Optical Range, Air Force Research Laboratory, 3550 Aberdeen Ave SE, Kirtland AFB, NM 87117-5776

David W. Tyler
Optical Sciences Center, The University of Arizona, 1630 E. University Blvd., Tucson, Arizona 85721
david.tyler@optics.arizona.edu

Abstract: Superresolution by data inversion is the extrapolation of measured Fourier data to regions outside the measurement bandwidth using postprocessing techniques. Here we characterize superresolution by data inversion for objects with finite support using the twin concepts of primary and secondary superresolution, where primary superresolution is the essentially unbiased portion of the superresolved data and secondary superresolution is the remainder. We show that this partition of superresolution into primary and secondary components can be used to explain why some researchers believe that meaningful superresolution is achievable with realistic signal-to-noise ratios, and other researchers do not.

References and Links

1. M. Bertero and C. De Mol, "Super-resolution by data inversion," in *Progress in Optics XXXVI*, E. Wolf, ed. (Elsevier, Amsterdam, 1996), pp. 129-178.
2. S. Bhattacharjee and M. K. Sundareshan, "Mathematical extrapolation of image spectrum for constraint-set design and set-theoretic superresolution," *J. Opt. Soc. Am. A* **20**, 1516-1527 (2003).
3. B. R. Hunt, "Super-resolution of images: algorithms, principles, and performance," *Int. J. Imaging Syst. Technol.* **6**, 297-304 (1995).
4. H. Liu, Y. Yan, Q. Tan, and G. Jin, "Theories for the design of diffractive superresolution elements and limits of optical superresolution" *J. Opt. Soc. Am. A* **19**, 2185-2193 (2002).
5. V. F. Canales, D. M. de Juana, and M. P. Cagigal, "Superresolution in compensated telescopes," *Opt. Lett.* **29**, 935-937 (2004).
6. C. K. Rushforth and R. W. Harris, "Restoration, resolution, and noise," *J. Opt. Soc. Am.* **58**, 539-545 (1968).
7. J. J. Green and B. R. Hunt, "Improved restoration of space object imagery," *J. Opt. Soc. Am. A* **16**, 2859-2865 (1999).
8. B. R. Frieden, "Evaluation, design, and extrapolation methods for optical signals based on the use of prolate functions," in *Progress in Optics IX*, E. Wolf, ed. (North-Holland, Amsterdam, 1971), pp. 313-407.
9. D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – I," *Bell System Tech. J.* **40**, 43-63 (1961).
10. H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – II" *Bell System Tech. J.* **40**, 65-84 (1961).
11. M. C. Roggemann and B. Welsh, *Imaging Through Turbulence* (CRC Press, Boca Raton, 1996).
12. W. P. Latham and M. L. Tilton, "Calculation of prolate functions for optical analysis," *Appl. Opt.* **26**, 2653-2658 (1987).
13. B. Porat, *Digital Processing of Random Signals, Theory and Methods* (Prentice-Hall, Englewood Cliffs, 1994), pp. 65-67.
14. R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Addison-Wesley, Reading, 1992), pp. 253-305.
15. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in FORTRAN*, 2nd ed., (Cambridge Press, Cambridge, 1996), pp. 134-135.
16. A. O. Hero, J. A. Fessler, and M. Usman, "Exploring estimator bias-variance tradeoffs using the uniform CR bound," *IEEE Trans. Sig. Proc.* **44**, 2026-2041 (1996).
17. P. Stoica and T. L. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Trans. Sig. Proc.* **49**, 87-90 (2001).

18. H. J. Landau and H. O. Pollack, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – III: the dimension of the space of essentially time-and band-limited signals," *Bell System Tech. J.* **41**, 1295-1336 (1962).

19. Y. L. Kosarev, "On the superresolution limit in signal reconstruction," *Soviet Journal of Communications Technology & Electronics* **35**, 90-108 (1990).

1. Introduction

Superresolution research has been conducted for many decades, both to develop ways to achieve superresolution and to characterize the properties of superresolution (see Refs.[1], [2], [3], and the references therein). In the research literature, the term "superresolution" is commonly used to describe the reduction in the point spread function (PSF) width in image-domain intensities as a result of applying some sort of processing methodology when compared to the PSF width in the corresponding unprocessed intensities. This definition results from the classical concept of resolution as proposed by Lord Rayleigh in which the resolution of a system is related to the minimum separation of two points that can be resolved. PSF width reduction can occur as the result of two different phenomena, either separately or in combination. The first phenomena is an increase in the Fourier amplitudes of the image-domain intensities within the measurement bandwidth. The Fourier amplitudes can be increased by postprocessing techniques or by including appropriate phase and/or amplitude pupil-plane filters in the optical system. This latter approach is often referred to as "optical superresolution" and is useful when a narrow pre-detection PSF is desired as is the case, for example, for optical disk readout [4] or for real-time resolution increases in adaptive optics data [5]. The second phenomena is the extrapolation of the measured Fourier data outside of the measurement bandwidth. By its very nature, only postprocessing can accomplish this second type of superresolution, and thus the term "superresolution by data inversion" has been coined to describe this type of superresolution [1]. The superresolution we explore in this paper is superresolution by data inversion.

There are two schools of thought on superresolution by data inversion, hereafter referred to simply as superresolution. People who subscribe to the first school of thought believe that meaningful superresolution (defined as accurate extrapolation of the measured data more than an incremental amount beyond the measurement bandwidth for measured data with realistic signal-to-noise ratios (SNRs) and space-bandwidth products (SBPs) $\gtrsim 10$) is unachievable and base their belief on signal-to-noise analyses carried out under the assumption that the object of interest has finite support in the image domain [1], [6]. The definition of SBP will be given in Section 2. The support of an object is the region in the image domain outside of which it is known that the object's intensities are zero. People who subscribe to the second school of thought believe that meaningful superresolution is possible for measured data with realistic SNRs and SBPs $\gtrsim 10$ and base their belief on the appearance of structure in the Fourier spectra of postprocessed imagery outside of the measurement bandwidth that appears to be correlated with the Fourier structure of the true object [3], [7].

In this paper, we describe a Fourier-domain decomposition of superresolved data into primary and secondary components in order to address the issue of whether or not meaningful superresolution for realistic SNRs is achievable. We define primary superresolution as that portion of the superresolved data that can be estimated in an essentially unbiased manner given the data SNRs, and secondary superresolution as the remainder of the superresolved data. The use of the phrase "essentially unbiased" will be explained and quantified in Section 4. We show that the first school of thought is based upon the ability to achieve primary superresolution and a component of secondary superresolution, while the second school of thought is based upon the ability to achieve secondary superresolution. The analysis throughout this paper is carried out in the context of achieving superresolution in the estimate of a 1-D support-constrained object, where the measured data is contained in a low-pass region of the Fourier domain. This problem formulation is particularly convenient because the inverse problem can be expressed analytically. In Section 2, we look at the forward and inverse problems in terms of their eigenvector/eigenvalue expansions and characterize the

inverse problem based upon the properties of the eigenvectors and eigenvalues. In Section 3, we analyze superresolution using the algorithm-independent Cramér-Rao lower bound (CRB) approach. The CRB analysis indicates that any unbiased estimates of Fourier data outside the measurement bandwidth have infinite variances. For this reason, in Section 4, the superresolution analysis is approached in terms of the eigenvector/eigenvalue expansion of the inverse problem where only a finite number of terms are kept in order to keep the estimator variances finite. The bias properties of the superresolved data are then explored as a function of the amount of extrapolation and the number of terms included in the expansion. From this analysis comes the decomposition of superresolved Fourier data into primary and secondary components. Additionally, in Section 4, properties of primary superresolution are described. In Section 5, properties of secondary superresolution are given. Finally, conclusions and future work are presented in Section 6.

2. Forward model

In this section, equations for the forward and inverse problems are given. Particular attention is placed upon the eigenvector/eigenvalue expansion of the inverse problem in the Fourier domain. Properties of these eigenfunctions and eigenvalues are presented in order to gain insight into the properties of the variances of unbiased estimates of the true object's Fourier data. An expression is given for these variances in terms of an infinite summation involving the eigenfunctions and eigenvalues and it is shown that the variances of the superresolved data appear to be infinite. Due to the lack of closed-form expressions for the eigenvectors, however, conclusive statements are not possible. This motivates the CRB approach developed in Section 3.

The forward model describing the functional dependence of the image i(x) on the noise-free object o(x), the system PSF h(x), and the additive noise n(x), is given by

$$i(x) = \int_{-x_0}^{x_0} h(x - \alpha)o(\alpha)d\alpha + n(x),$$
(1)

where h(x) is the inverse Fourier transform of an ideal low-pass filter with a cutoff frequency of f_0 (in cycles/unit length) and n(x) is stationary zero-mean Gaussian noise with variance σ [6], [8]. It is straightforward to generalize this forward model to include other PSFs and noise models, but this simple and oft-used image model is chosen for purposes of clarity. The eigenfunctions $\{\tilde{\phi}_m(x)\}$, m=0,1,2,... , of this forward model are the celebrated prolate spheroidal wave functions (PSWFs) of Slepian, Pollack, and Landau [9], [10], and are a function of the space-bandwidth product (SBP) $4x_0f_0$ [8]. Using these PSWFs and their associated eigenvalues $\tilde{\lambda}_m$, Eq.(1) can be rewritten as

$$i(x) = \sum_{m=0}^{\infty} \tilde{\lambda}_m \tilde{\phi}_m(x) \int_{-x_0}^{x_0} o(\xi) \tilde{\phi}_m(\xi)d\xi + n(x),$$
(2)

where the PSWFs are normalized to unit energy in the interval [-∞,∞] and the eigenvalues are bounded above by one. Because the object has finite support, the eigenvalues are greater than zero for all m, but their values are a decreasing function of m and converge to zero (in fact, the eigenvalues are the energies of their associated PSWFs in the object's support). For this reason, the inverse of Eq.(2) exists but is discontinuous.

Since our interest is in the properties of estimates of the object's Fourier spectrum outside the measurement bandwidth, we formulate the inverse problem in terms of estimating the true object spectrum O(f) from the measured Fourier spectrum I(f) = O(f) + N(f). The desired expression is given by

$$O_e(f) = \sum_{m=0}^{M} \lambda_m^{-1} \varphi_m(f) \int_{-f_o}^{f_o} O(\xi) \varphi_m(\xi) d\xi + \sum_{m=0}^{M} \lambda_m^{-1} \varphi_m(f) \int_{-f_o}^{f_o} N(\xi) \varphi_m(\xi) d\xi, \qquad (3)$$

where $N(f)$ is the Fourier transform of $n(x)$ and $O_e(f)$ is the estimate of the object spectrum. The PSWFs of Eq.(3) are scaled versions of the PSWFs in Eq.(2) and their associated eigenvalues are the energies of the PSWFs inside the measurement bandwidth [9]. Because $\varphi_m(f)$ has infinite support for each value of m, the summations in Eq.(3) must be over all values of m in order to obtain an unbiased estimate of $O_e(f)$ at any frequency. Although any practical application can include only a finite number of terms (M finite), analyzing the infinite-summation case is important because it provides insight into the SNR properties of unbiased estimates (M=∞). From Eq.(3), it is easily seen that the expected value of $O_e(f)$ is just the true object spectrum. To obtain the SNRs of $O_e(f)$ at any frequency, an expression for the variances of $O_e(f)$ are needed and can be obtained in a straightforward manner from Eq.(3), yielding

$$\text{var}\{O_e(f)\} \overset{\Delta}{=} \text{E}[O_e(f)O_e^*(f)] - |\text{E}[O_e(f)O_e^*(f)]|^2$$
$$= \sigma^2 \sum_{m=0}^{\infty} \lambda_m^{-1} \varphi_m^2(f) \qquad (4)$$

The standard imaging definition of SNR is the expected value of the estimated quantity divided by the square root of its variance [11]. It can be seen from Eqs.(3) and (4) that the SNRs of $O_e(f)$ depend upon the properties of the PSWFs and their eigenvalues, and is a function of the inverse of the sum of the square roots of the PSWF eigenvalues. For these reasons, these properties are explored next.

In general, PSWFs for any value of the SBP have similar properties when appropriately normalized. For example, any PSWF whose index satisfies m≤SBP has most of its energy inside the measurement bandwidth. Most of the energy of any PSWF whose index satisfies m≥SBP is outside the measurement bandwidth. We call every PSWF whose index satisfies m≥SBP a *superresolving PSWF* for this reason. To gain insight into additional properties of PSWFs, it is useful to look at a representative set of PSWFs as a function of frequency and index value. Using the method of Latham and Tilton [12], for any SBP we are able to generate PSWFs and their eigenvalues quickly and accurately down to $\lambda_m \approx 10^{-30}$.

Fig. 1 is a movie displaying PSWFs as a function of their index for a SBP of 40. These PSWFs were generated in a 1024 element array of which 256 elements were in the measurement bandwidth. Their associated eigenvalues are plotted in Fig. 2. In addition to the energy properties described in the previous paragraph, notice that the superresolving PSWFs become essentially decaying sinusoids, whose starting location f_m in frequency space monotonically increases as a function of m. This implies that the energies of the superresolving PSWFs become increasingly concentrated at higher and higher frequencies. Recall, however, that the energy of any PSWF inside the measurement bandwidth is equal to its eigenvalue; thus, from Fig. 2., it is clear that every PSWF has non-zero energy inside the measurement bandwidth. In fact, this property is the foundation for superresolution in the sense that there does not exist any finite index value M for which the inner product shown in the first summation of Eq.(3) is zero. As a result, noise-free knowledge of the true object's Fourier spectrum inside the measurement bandwidth permits superresolving this knowledge to all frequencies.

The primary difference in the properties of PSWFs and their eigenvalues as the SBP is varied is the decay rate of the eigenvalues. When this difference impacts the properties of primary and secondary superresolution, it will be called out. Otherwise, results described here apply to PSWFs for any SBP value.
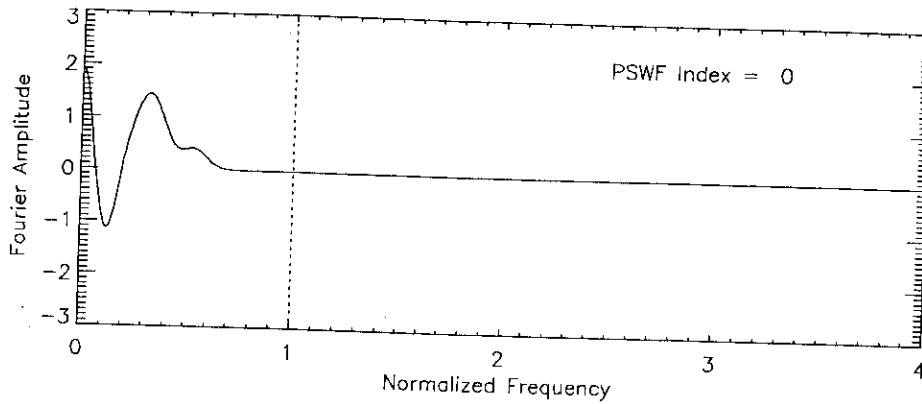
Fig. 1. A movie of PSWF Fourier amplitudes plotted as a function of frequency for a SBP of 40. The frequency axis is normalized to one at the edge of the measurement bandwidth. Each movie frame corresponds to a PSWF for the PSWF index indicated in the movie.

Returning now to the analysis of the SNRs of $O_e(f)$ for all frequencies, recall that the expected value of $O_e(f)$ is just the true object spectrum. Therefore, for the SNRs to remain finite, var$\{O_e(f)\}$ must remain finite. From Eq.(4), it can be seen that the question of whether or not var$\{O_e(f)\}$ is finite for any frequency is determined by the index dependence of the ratio $\varphi_m^2(f)/\lambda_m$. From Fig. 1, it can be seen that $\varphi_m^2(f) \to 0$ for any frequency f as $m \to \infty$, and from Fig. 2 that $\lambda_m \to 0$ as $m \to \infty$. As a result, their relative rates of convergence determine whether or not var$\{O_e(f)\}$ remains finite. Because there does not exist an analytical expression for $\varphi_m(f)$, Eq.(4) cannot be used to definitively answer this question. However, we can use the PSWFs and eigenvalues calculated for Figs. 1 and 2 to calculate var$\{O_e(f)\}$ for values of λ_m down to $\approx 10^{-30}$ to see if it appears to be converging to a finite number for any frequency. A movie of var$\{O_e(f)\}$ as a function of the upper limit of the summation in Eq.(4) is shown in Fig. 3. A key property seen in this movie is that var$\{O_e(f)\}$ does not seem to be converging for any frequency, even those inside the measurement bandwidth, although the rate of increase of var$\{O_e(f)\}$ inside the measurement bandwidth is significantly less than outside. Because of this fact, it appears that the estimator given by Eq.(3) can only provide an unbiased estimate of any frequency of a support-constrained object at the cost of an infinite variance.

The conclusion that an unbiased estimate of a support-constrained object's Fourier spectrum at any frequency has infinite variance was based upon a specific estimator (i.e., Eq.(3)). It is natural to wonder if an alternate estimator exists that can produce an unbiased estimate with finite variance for at least some frequencies. To generate an algorithm-independent answer, the next section uses a CRB approach to calculate the variances of any estimator's unbiased estimate of a support-constrained object's Fourier spectrum.
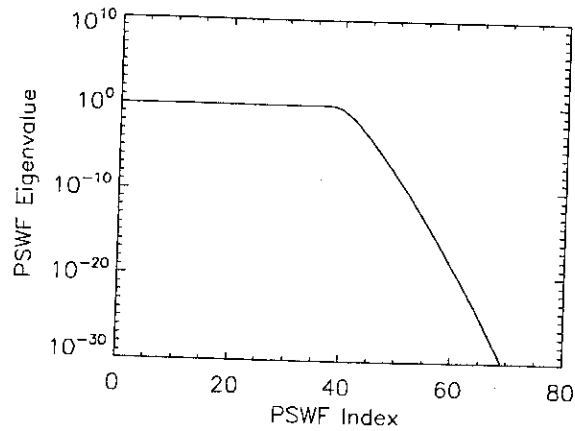
5

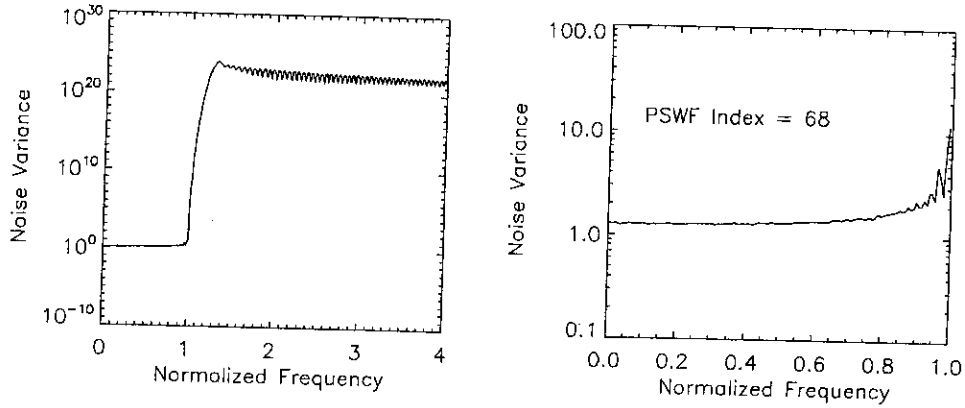Fig. 2. A plot of the PSWF eigenvalues associated with the PSWFs in Fig. 1.

Fig. 3. A movie of the Fourier variances defined by Eq.(4) for a SBP of 40 plotted as a function of frequency normalized to one at the edge of the measurement bandwidth. Each movie frame corresponds to increasing the finite summation limit in Eq.(4) by one. The plots on the left are for all frequencies included in the calculations, while the plots on the right are a magnified version of the plots on the left for frequencies within the measurement bandwidth.

3. Cramér-Rao Bounds

In this section, an algorithm-independent approach for calculating a particular set of lower bounds to the variances of any unbiased estimate of $O(f)$ is employed. These lower bounds, called CRBs [13], are the diagonal elements of the inverse of the Fisher information matrix (FIM) corresponding to $O(f)$ and the conditional measurement probability density function (PDF) that is parameterized by $O(f)$. Our approach to calculating these CRBs is to first generate the FIM corresponding to $o(x)$, then take its inverse, then transform this inverse into the inverse of the FIM corresponding to $O(f)$, and then extract its diagonal elements. Although the CRBs could have been calculated by creating the FIM for $O(f)$ directly, it will be shown that the transformation process from the image-domain inverse FIM to the Fourier-domain inverse FIM provides important insight into the properties of the Fourier-domain CRBs.

The unbiased CRBs corresponding to a vector of parameters are lower bounds to the variances of any unbiased estimates of these parameters. CRBs for a function of these

parameters can be calculated when the Jacobian of the function with respect to the parameters is known. Although CRBs are not necessarily achievable, in many cases they can be achieved and in other cases they can be approached closely. To generate the CRBs for a vector of parameters, the PDF of the uncertainty in the measurements, conditioned on the parameter values, must be known. For the imaging model of Eq.(1), this conditional PDF can be obtained from the noise PDF and the noise-free imaging model. However, because CRB theory applies to a vector of random variables, not a stochastic process, Eq.(1) must be rewritten in a sampled data format. For this reason, we replace the continuous variable x by a set of evenly-spaced locations covering the interval $[-x_o, x_o]$ for some arbitrary number of locations. Let α be a vector containing these locations, and let y, θ, and η be vectors that contain the values of $i(\alpha)$, $o(\alpha)$, and $n(\alpha)$, respectively. In addition, let H be the circulant matrix associated with $h(\alpha)$ [14]. This permits rewriting Eq.(1) as a matrix-vector equation given by

$$y = H\theta + \eta,$$

(5)

where the PDF of y conditioned on θ is given by

$$f(y; \theta) = f_\eta(y - H\theta),$$

(6)

and $f_\eta(\eta - H\theta)$ is the joint pdf of η but with mean $H\theta$.

As mentioned earlier in this section, the first step in calculating the CRBs for any unbiased estimate of $O(f)$ is to calculate the FIM corresponding to θ, $F(\theta)$. The element of $F(\theta)$ in the pth row and the qth column is given by

$$F_{p,q}(\theta) = E\left[\frac{\partial \ln f(y; \theta)}{\partial \theta_p} \frac{\partial \ln f(y; \theta)}{\partial \theta_q}\right],$$

(7)

where ln denotes the natural logarithm and E[] is the expected value of the expression in the brackets. For the imaging model of Eq.(5), Eq.(7) becomes

$$F_{p,q}(o) = \frac{1}{\sigma^2} \sum_k \left\{\frac{\partial(H\theta)_k}{\partial \theta_p} \frac{\partial(H\theta)_k}{\partial \theta_q}\right\}$$

$$= \frac{1}{\sigma^2} \sum_k h(\alpha_k - \alpha_p) h(\alpha_k - \alpha_p)$$

(8)

where the second equality in Eq.(8) is written in terms of h and the vector of x locations α rather than in terms of the elements of H to enhance the reader's understanding of how we calculate $F(\theta)$. Notice that each element of $F(\theta)$ is the inverse of the noise variance multiplying a discrete approximation to the integral of the product of two shifted versions of the PSF. This discrete integration approximation must be accurate to greater than 30 orders of magnitude in order to generate results comparable in accuracy to the noise variance calculations in Section 2. One approach to obtaining this accuracy is to densely sample the support of the object; however, that leads to large dimensions for $F(\theta)$ and prohibitively long calculations. A second approach, and the one we implemented, is to carry out the full integration implied by the summation in Eq.(8). We used a quad-precision Romberg integration routine based upon the Numerical Recipes' qromb routine [15].

The second step is to calculate the inverse of $F(\theta)$. Prior to calculating this inverse for each case of interest, the eigenvalues were obtained to determine the stability of the inverse. A set of eigenvalues for the same set of parameters as the results in Section 2 (SBP of 40, vector size of 1024 elements, and an object support size of 161 elements) are plotted in Fig. 4 as a function of the FIM eigenvalue index. Several important issues are apparent from this

(6)

figure. First, the FIM eigenvalue behavior as a function of the FIM eigenvalue index is quite similar to the PSWF eigenvalue behavior as a function of the PSWF index in that the eigenvalues are close to one for values of the FIM eigenvalue index $\leq$ SBP and decay at approximately the same rate as do the PSWF eigenvalues. Second, because of the similarity of the behaviors and values of the PSWF eigenvalues and the FIM eigenvalues, it is reasonable to assume that the FIM eigenvalues converge to zero. For this reason, a true inverse of $F(\theta)$ cannot be calculated. Fortunately, unbiased CRBs can be calculated using a pseudoinverse of $F(\theta)$ [16]. However, the fact that the pseudoinverse must be used implies that unbiased estimates with finite variances of all the parameters in θ are not possible [17]. As a result, only functions of θ have the possibility of being estimated in an unbiased manner with finite variances. Fortunately, our interest is in functions of θ; specifically, we are interested in the Fourier transform of θ. This is discussed next.
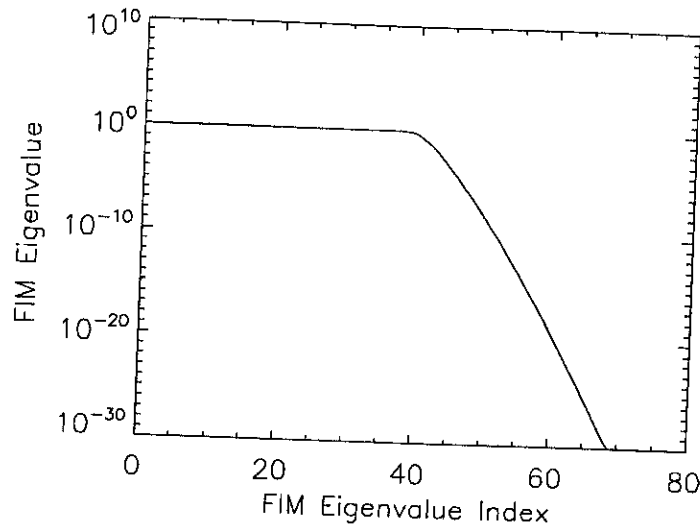

Fig. 4. A plot of the FIM eigenvalues.

The third step is to transform the pseudoinverse of $F(\theta)$, $F(\theta)^\dagger$, into the inverse of the FIM corresponding to the Fourier transform of θ and extract the CRBs that are along its diagonal. The desired transformation is brought about by premultiplying $F(\theta)^\dagger$ by the Jacobian of the Fourier transform operation G and postmultiplying by G^T. If the perpendicular to the null space of $F(\theta)^\dagger$ is contained in the range space of G^T for some finite set of FIM eigenvalues included in the pseudoinverse, then the CRBs of the Fourier transform of θ are finite [17]. More specifically, if the Fourier transform of θ at a given frequency has no component in the null space of $F(\theta)^\dagger$ for some finite set of FIM eigenvalues included in the pseudoinverse, then an unbiased estimate of the Fourier transform of θ with finite variance at that frequency is possible. What this means in practice is that, if an unbiased estimate of the Fourier transform of θ at a given frequency is possible with finite variance, then it can be identified by the fact that its CRB as a function of the number of eigenvalues included in $F(\theta)^\dagger$ converges to a finite value as the number of FIM eigenvalues included in $F(\theta)^\dagger$ increases.

Based upon the discussion in the previous paragraphs, the unbiased CRBs for estimating $O(f)$ as a function of the number of FIM eigenvalues included in the calculation of $F(\theta)^\dagger$ are displayed in the movie in Fig. 5 for the same set of parameters as used for generating Fig. 4.

Each frame of the movie corresponds to adding one more eigenvalue to the calculation of $F(\theta)^\dagger$ than in the previous frame. Two important observations can be made from this movie. The first is that none of the CRBs appear to be converging to a finite value. Even the CRBs inside the measurement bandwidth do not appear to be converging. This slow increase in the CRBs inside the measurement bandwidth is undoubtedly due to the "leakage" of the large CRB values outside the measurement bandwidth into the measurement bandwidth brought about by the Fourier-domain correlations enforced by the support constraint. The CRBs outside the measurement bandwidth are increasing rapidly as a function of the FIM eigenvalues included in the calculation of $F(\theta)^\dagger$. The second observation is that the CRB plots in Fig. 5 are remarkably similar to the noise variance plots of Fig. 3.
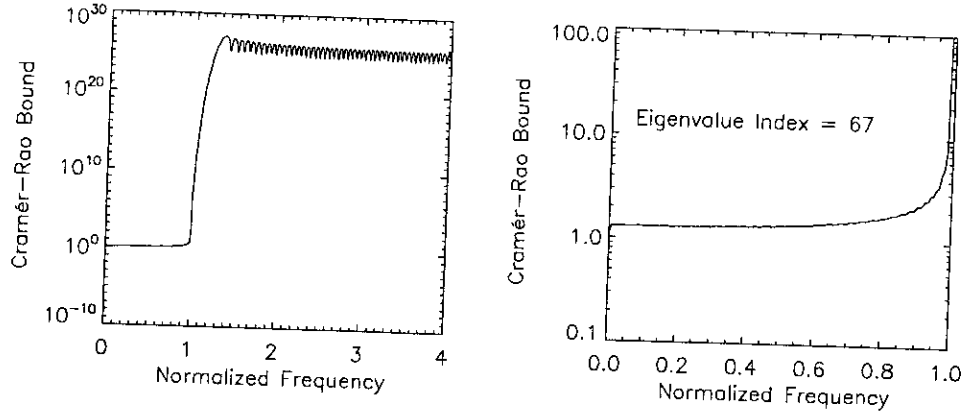

Fig. 5. A movie of the CRBs for any unbiased estimate of any Fourier spectra for a SBP of 40 plotted as a function of frequency normalized to one at the edge of the measurement bandwidth. Each movie frame corresponds to adding one additional eigenvalue to the pseudoinverse of the FIM. The plots on the left are for all frequencies included in the calculations, while the plots on the right are a magnified version of the plots on the left for frequencies within the measurement bandwidth. These plots are independent of the actual Fourier spectra under consideration because the measurement noise is signal independent.

Based upon the preceding CRB analysis as well as the results in Section 2, it seems clear that no unbiased estimate of O(f) for any superresolved frequency is possible with finite variance. Because values of O(f) inside the measurement bandwidth are available from the measurement directly, they can be estimated in an unbiased manner with finite variances merely by taking as their estimates the measurement values themselves. However, no unbiased superresolution with finite variance is possible. For this reason, biased estimators are explored in the next section, leading to the concepts of primary and secondary superresolution.

4. Primary superresolution

In this section, it is shown that, although unbiased superresolution with finite variances is not achievable, essentially unbiased superresolution with finite variances is achievable. Because biases are a function of the estimator, first a PSWF-based estimator is proposed and its use justified. Next, a quantitative definition of the phrase "essentially unbiased" is derived, and superresolved data is characterized as primary superresolution if it is essentially unbiased. Finally, some properties of primary superresolution are presented.

To explore biased superresolution, an estimator must be specified. We choose to use a finite-summation version of the PSWF-based inverse problem formulation in Eq.(3) for three reasons. The first is that, for any finite upper limit M to the sum in Eq.(3), the PSWFs have the most energy in the measurement bandwidth of any basis set [18]. The benefit this brings to the inverse problem can be seen from Eq.(4), where it is shown that the noise variances are

inversely proportional to the PSWF eigenvalues. Because these eigenvalues are the energies of the PSWFs inside the measurement bandwidth, the noise levels in the superresolved data for any set of M basis functions are minimized by using PSWFs as the basis set. The second reason is that the similarities between the pseudo-inverse-based CRBs shown in Fig. 5 and the variances for this estimator shown in Fig. 3 imply that the PSWF-based estimator is in some sense qualitatively optimal. The third reason is that a closed-form solution to the inverse problem permits straightforward analysis of the bias properties of superresolution.

To begin the analysis, consider Eq.(3). Clearly the bias properties of this (and any other) estimator are a function of the object being estimated. However, if the biases in the superresolved data produced by this estimator could be characterized in a way that is independent of the object being estimated, the result would be more general and thus more useful. Such a characterization is possible and is derived next based upon the frequency behavior of the PSWFs. Because the estimator in Eq.(3) generates an estimate of O(f) using linear combinations of PSWFs, it follows that PSWFs with minimal energy at a given frequency contribute little to $O_e(f)$ at that frequency. In addition, the movie of the PSWFs in Fig. 1 shows that the energies of superresolving PSWFs are close to zero in a connected region outside of but adjacent to the measurement bandwidth, and that the size of this region is an increasing function of the PSWF index. The combination of these two facts implies that the bias at a given superresolved frequency is a function of the energies at that frequency contained in PSWFs not included in the summation in Eq.(3) and that this bias is a decreasing function of the summation upper limit. At a given superresolved frequency, if the energies of the PSWFs not included in the estimator are negligible, the superresolved data can be said to be essentially unbiased.

To quantify this discussion, the sum of the energies of the PSWFs not included in the inverse problem must be calculated as a function of superresolved frequency and PSWF index upper limit. To present the results with increased clarity, these sums were divided by the total energies of the PSWFs at each superresolved frequency to show the fraction of energy not included. The results of these calculations are shown in the movie in Fig. 6, where each frame of the movie is for a given superresolved frequency. The PSWFs used for this movie are the PSWFs used to generate the movie in Fig. 1, but the general shapes of the curves in Fig. 6 are similar for all SBPs. The impact resulting from the fact that the PSWFs are not identical for all SBPs will be discussed later in this section. Notice that the PSWF energy not included in the inverse problem for a given superresolved frequency decays rapidly as a function of PSWF index and asymptotes to zero. The curves clearly show a "knee in the curve" near the zero point that is arrived at when approximately 2% of the PSWF energy is not included. For this reason, we call the superresolved data essentially unbiased when more than 98% of the total PSWF energy is included in the estimation of the object's spectrum at that point, and thus have quantified the phrase "essentially unbiased". Although the 98% point is somewhat arbitrary, the conclusions to be drawn from this definition of primary superresolution are relatively insensitive to the exact percentage used in the definition.
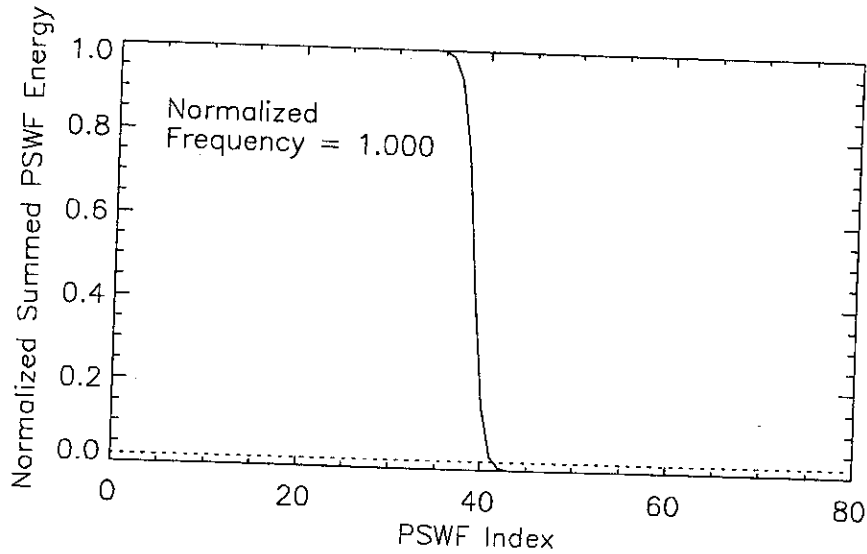
Fig. 6. A movie of the normalized residual energy of the PSWFs not included in the reconstruction process as a function of the PSWF index for a SBP of 40. The dotted line corresponds to a value of 0.02. Each frame of the movie corresponds to a different superresolved frequency as indicated in the movie.

Some properties of primary superresolution will now be discussed. First, as noted previously, the region in Fourier space where primary superresolution is achieved is an increasing function of the maximum PSWF index value allowed in the inverse problem. Because the maximum index value chosen is a function of the data SNRs due to the noise amplification process described by the summation in Eq.(4), this indicates that the amount of primary superresolution is a function of the data SNRs. This conclusion is well known, of course, and shows that our definition of primary superresolution is consistent with previous results. Additional properties of primary superresolution can be seen by examining Fig. 7, where the amount of primary superresolution is plotted as a function of the inverse of the square root of the PSWF eigenvalues. As noted in the discussion following Eq.(4), the SNRs of the estimated Fourier spectrum are related to the inverse of the square root of the PSWF eigenvalues, so Fig. 7 shows the amount of primary superresolution possible as a function of the recovery SNR. The vertical axis of Fig. 7 is the amount of primary superresolution achieved, and is expressed in terms of the number of degrees of freedom added to the image reconstruction. In this context, a degree of freedom refers to the classical sampling degree of freedom that is defined as the inverse of two times the maximum frequency included in the image reconstruction [8].

Several additional properties of primary superresolution can be deduced from Fig. 7. The first is that, for SBPs $\gtrsim 10$, the amount of primary superresolution is essentially independent of the SBP. In other words, this means that primary superresolution is additive in the sense that, regardless of the object support size or the size of the Fourier-domain measurement bandwidth, as long as the SBP is not on the order of one, the amount of primary superresolution possible is independent of the SBP and adds degrees of freedom. The second observation is that the number of degrees of freedom added is a logarithmic function of the data SNR, a conclusion also reached previously by a number of researchers [1], [19]. The third observation is surprising, however. Notice that the amount of primary superresolution possible for poorly-resolved objects (i.e., for objects whose SBPs are on the order of one) is less than that possible for well-resolved objects. This conclusion apparently contradicts previous results [1]; however, the apparent contradiction is resolved by clarifying how

superresolution is measured. We measure superresolution in terms of the number of degrees of freedom added to the reconstructed object. If, however, superresolution is viewed in terms of the percentage increase in the number of degrees of freedom as is done in [1], the increase in resolution for poorly-resolved objects is much greater than for well-resolved objects because poorly-resolved objects start out with very little information content. We argue that the best way to view primary superresolution is in terms of adding degrees of freedom, rather than in terms of percentage increase in the number of degrees of freedom because superresolution is an inherently additive phenomenon and can be best understood using an additive approach. In the next section, where secondary superresolution is explored, it will be shown that an additive degrees-of-freedom approach is also the appropriate way to view superresolution; however, the definition of degree of freedom changes.
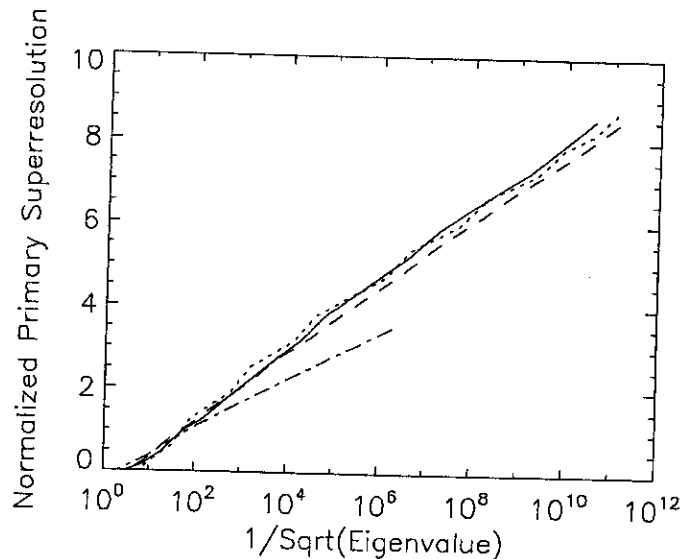

Fig. 7. A plot of the increases in the sampling degrees of freedom in a superresolved reconstruction brought about by including primary superresolution components. The dot-dash line is for a SBP of 1, the dashed line is for a SBP of 10 , the solid line is for a SBP of 20, and the dotted line is for a SBP of 40.

5. Secondary superresolution

The concept of secondary superresolution is discussed in this section and its properties are compared and contrasted with the properties of primary superresolution. The amount of superresolution possible when including both primary and secondary superresolved data is evaluated using a PSWF-based degrees-of-freedom approach. It is shown that secondary superresolution is a more-than-additive phenomena and that the amount of superresolution is a function of location in the object support, unlike primary superresolution.

Secondary superresolution is defined as all of the superresolved data that is not essentially unbiased. Another way to express this fact is to call secondary superresolution either biased or incomplete. For this reason, the classical degrees-of-freedom relationship relating the number of independent components in the image domain with the maximum frequency included in the reconstruction process no longer holds. To illustrate this fact, consider the superresolved Fourier spectra of a simulated 1-D triple star shown in Fig. 8. The triple star was created in a 1024 element vector and consists of three equal-magnitude point sources located at elements 437, 512, and 552. The true Fourier amplitudes are plotted along with the reconstructed Fourier amplitudes using Eq.(3) with an upper limit to the summation

of 46. The PSWFs from Fig. 1 were used for this reconstruction. Because these PSWFs were created using a SBP of 40, the upper summation limit of 46 means that six superresolving PSWFs were included in the reconstruction process. Notice that the region of primary superresolution is indeed essentially unbiased, while the region of secondary superresolution is strongly biased. Although the normalized frequency axis in Fig. 8 extends only to twice the measurement bandwidth, the superresolved Fourier spectra has infinite support because the structure of superresolving PSWFs is that of decaying sinusoids. Clearly infinite amounts of superresolution are not possible based upon the variance analyses of Sections 2 and 3. Thus the classical degree-of-freedom model is not valid to characterize the resolution properties of secondary superresolution. An alternate model must be used.
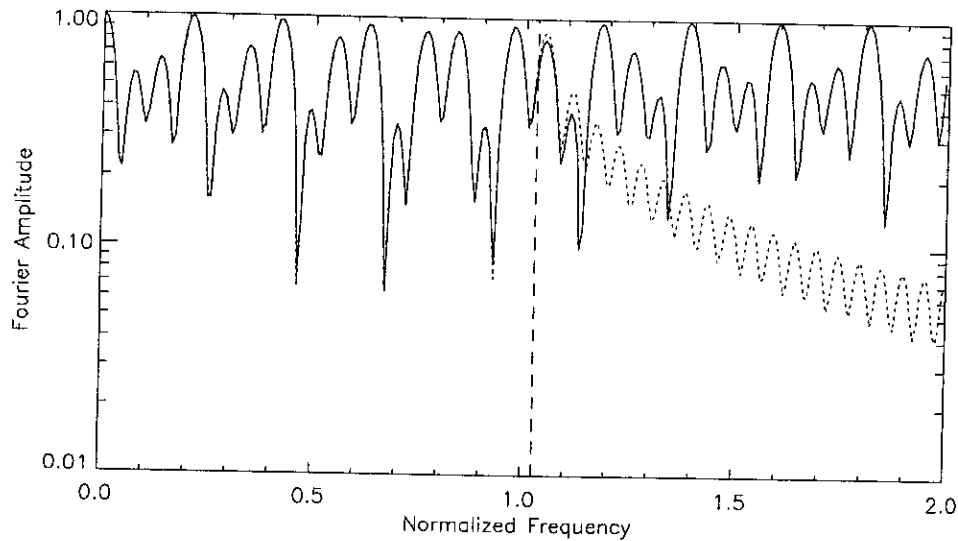

Fig. 8. Plots of the Fourier amplitudes of a triple star for the parameters listed in the text. The solid line is a plot of the true Fourier amplitudes while the dotted line corresponds to a reconstruction of the true Fourier amplitudes using Eq.(3) with an upper limit of 46 and a SBP of 40. The frequency axis is normalized to one at the edge of the measurement bandwidth, and the dashed line indicates the upper boundary of the region of primary superresolution.

It has been shown previously that a PSWF of index m has m zeros in the support region used to generate the PSWFs [9]. Based upon this fact, it has been proposed that the resolution provided by the mth PSWF can be characterized by the average distance between these zeros [6]. Such a definition of resolution agrees with the classical Rayleigh resolution limit when m is equal to the SBP. Based upon this definition of resolution, the number of degrees of freedom in a reconstructed image using PSWFs is equal to the number of PSWFs included in the reconstruction. This definition is consistent with the general concept of resolution in the classical sense and does not predict infinite resolution when including an infinite number of frequencies as is the case for the sampling degree of freedom.

Using the PSWF definition of degrees of freedom, the increase in the resolution of a reconstruction brought about by including both the primary and secondary superresolution components in the reconstruction process is plotted in Fig. 9 as a function of the inverse of the square root of the PSWF eigenvalue and for several values of the SBP. There are several properties of the increased degrees of freedom brought about by including secondary superresolution in the reconstruction process worth noting. The first is that poorly-resolved objects still benefit the least from superresolution process as measured by increased degrees of freedom. The second is that the amount of increase in degrees of freedom is no longer independent of the SBP for SBP $\gtrsim$ 10. In fact, the number of degrees of freedom in a

/3

superresolved reconstruction of an object is an increasing function of the number of degrees of freedom in the measurement. Thus the inclusion of secondary superresolution makes the superresolution process more than additive in terms of our definition of additive superresolution given in Section 4. For example, primary superresolution produces one additional degree of freedom when data SNR is 100 for all SBP values $\gtrsim$ 10, while the combination of primary and secondary superresolution produces four additional degrees of freedom when the SBP is 10, and five additional degrees of freedom when the SBP is 40. Thus it appears that secondary superresolution can noticeably increase the resolution in superresolved reconstructions as compared to including just primary superresolution.
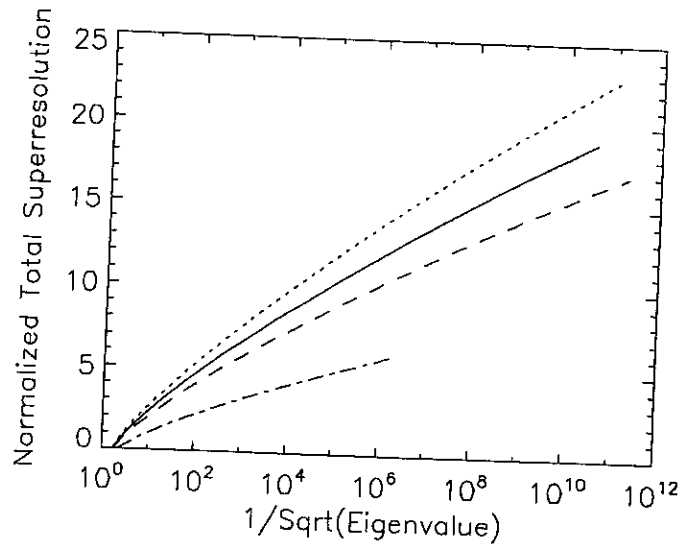

Fig. 9. Plots of the increases in the PSWF-defined degrees of freedom averaged over the entire support region brought about by including both primary and secondary superresolution in the reconstruction process. The dot-dash line is for a SBP of 1, the dashed line is for a SBP of 10, the solid line is for a SBP of 20, and the dotted line is for a SBP of 40.

However, defining the resolution of a given superresolving PSWF in terms of the average spacing of its zeros across the entire support region implies that the zeros are evenly spaced throughout the support region. As can be seen for a representative superresolving PSWF in Fig. 10, this is not the case. The zeros are relatively evenly spaced in the center of the support region, but their spacings decrease rapidly as one approaches the edge of the support. This implies that superresolving PSWFs produce higher resolutions near the edges of the support region and lower resolutions away from the edges of the support region than predicted by Fig. 9. To take the space-dependent resolution properties of superresolving PSWFs into account, the support region was divided into two regions, the center 75% and the remaining 25% at the edges. The amount of total superresolution achieved in the center 75% of the support region is plotted in Fig. 11 and can be seen to be significantly less than predicted by Fig. 9 for the entire support region. In order to compare directly the secondary superresolution results to the original degrees of freedom in the image, the results plotted in Fig. 11 correspond to the increase in the number of degrees of freedom that would occur in the entire support region if the resolution calculated using the center 75% of the support region is the same throughout the entire support region. In addition, notice that the amount of total superresolution is essentially independent of the SBP for SBP $\gtrsim$ 10, just as is the case for primary superresolution, and thus is additive by our definition. As before, the amount of total superresolution achievable is less for SBP values on the order of one. A difference from the primary superresolution results is

that the amount of superresolution is a little larger when including secondary superresolution in the reconstruction process.
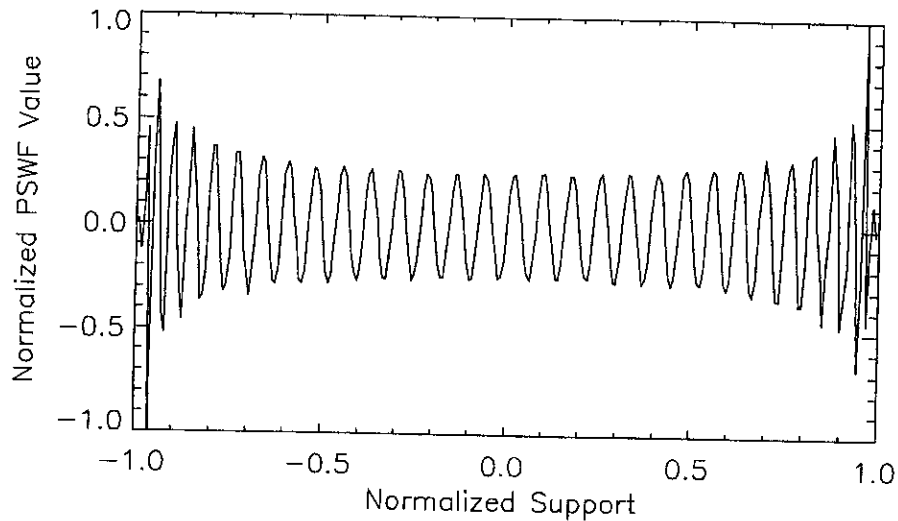

Fig. 10. Plot of the PSWF with index 65 for a SBP=40 demonstrating that the spacings of zeros for superresolving PSWFs in the support region decrease as the location in the support gets close to the support boundaries. Notice that the zeros are relatively evenly spaced in the center 75% of the object support.
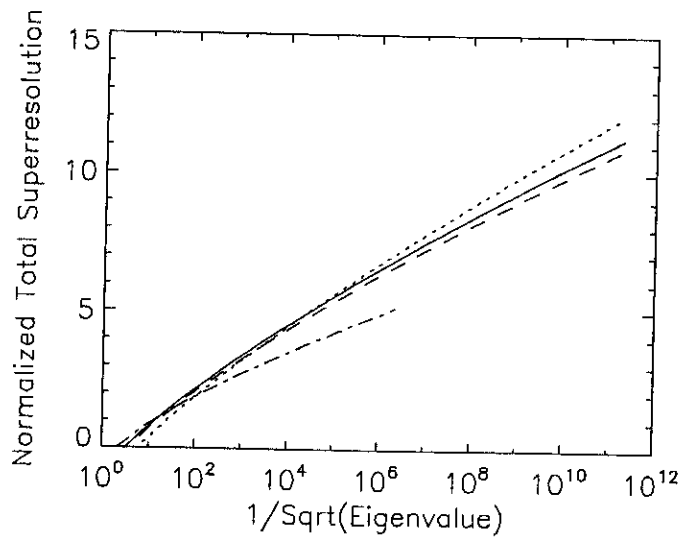

Fig. 11. Plots of the increases in the PSWF-defined degrees of freedom averaged over the center 75% of the support region brought about by including both primary and secondary superresolution in the reconstruction process. The dot-dash line is for a SBP of 1, the dashed line is for a SBP of 10, the solid line is for a SBP of 20, and the dotted line is for a SBP of 40. The vertical axis is the number of degrees of freedom that would be added to the entire support region if the resolution increase calculated in the center of the support region is the same throughout the support region.

The amount of total superresolution achieved at the edges of the support region for the remaining 25% of the support region is shown in Fig. 12. Notice the large increase in the numbers of degrees of freedom at the edges of the support. As for the results in Fig. 11, the numbers in Fig. 12 correspond to the increase in the degrees of freedom that would occur in the entire support region if the resolution at the edges of the support region is the same throughout the support region.
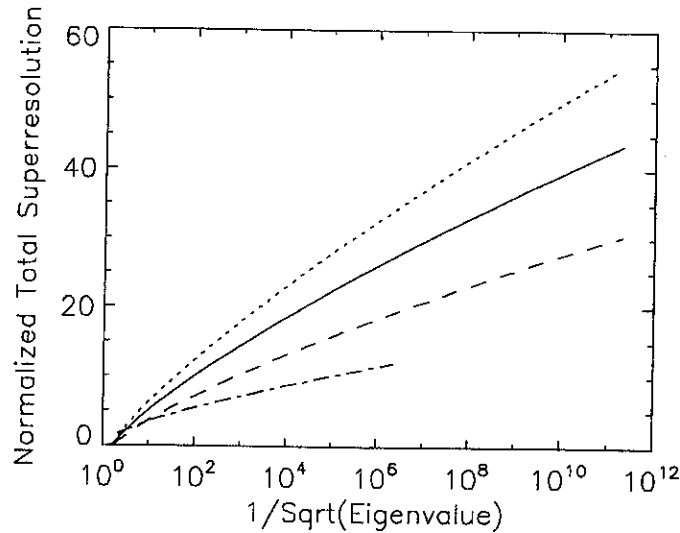

Fig. 12. Plots of the increases in the PSWF-defined degrees of freedom averaged over the 25% of the support region located at its edges brought about by including both primary and secondary superresolution in the reconstruction process. The dot-dash line is for a SBP of 1, the dashed line is for a SBP of 10 , the solid line is for a SBP of 20, and the dotted line is for a SBP of 40. The vertical axis is the number of degrees of freedom that would be added to the entire support region if the resolution increase calculated at the edges of the support region is the same throughout the support region.

To put the primary and secondary superresolution results in perspective, it is useful to consider the relative percentage increases in resolution for a specific case. To this end, consider an image that has a SBP of 40. If the data SNR allows the use of all the PSWFs whose eigenvalues are greater than 10^{-4}, the percent increase in the resolution throughout the entire support region due to primary superresolution would be 2.5% because primary superresolution only provides an additional degree of freedom. The resolution increase including both primary and secondary superresolution in the center region of the support would be 5%, a factor of two increase over what primary superresolution provides, but still small. Finally, at the edges of the support region, the resolution increase including both primary and secondary superresolution would be 30%, a large and noticeable increase. To demonstrate visually the impact of these numbers, three reconstructions of the triple star whose Fourier magnitude spectrum is plotted in Fig. 8 are displayed in Fig. 13. The first reconstruction uses just the Fourier data inside the measurement bandwidth, the second reconstruction uses both the measured Fourier data and the primary superresolution data, and the third reconstruction uses all the superresolved data along with the measured Fourier data. The two rightmost components of the triple star are contained inside the center 75% of the support; as a result, the amount of superresolution possible is predicted to be 5% for all three reconstructions and thus essentially negligible. The reconstructions in Fig. 13 confirm this

prediction. However, the leftmost component of the triple star is near the edge of the support region and benefits from the increased resolution brought about by secondary superresolution. As predicted by the plots in Fig. 7, the primary superresolution reconstruction of this component differs very little from the no-superresolution reconstruction. However, as predicted by the plots in Fig. 12, the reconstruction that uses secondary superresolution data has an increased resolution of ~30%. The increase in resolution is greater on the side of the component near the edge of the support as compared to the side nearer the center of the support because of the decreasing zero spacings of PSWFs as one moves toward the support boundaries. Although the increased resolution is encouraging, it can be seen that the reconstruction using secondary superresolution data produces less accurate estimates of the relative magnitudes of the triple star components.
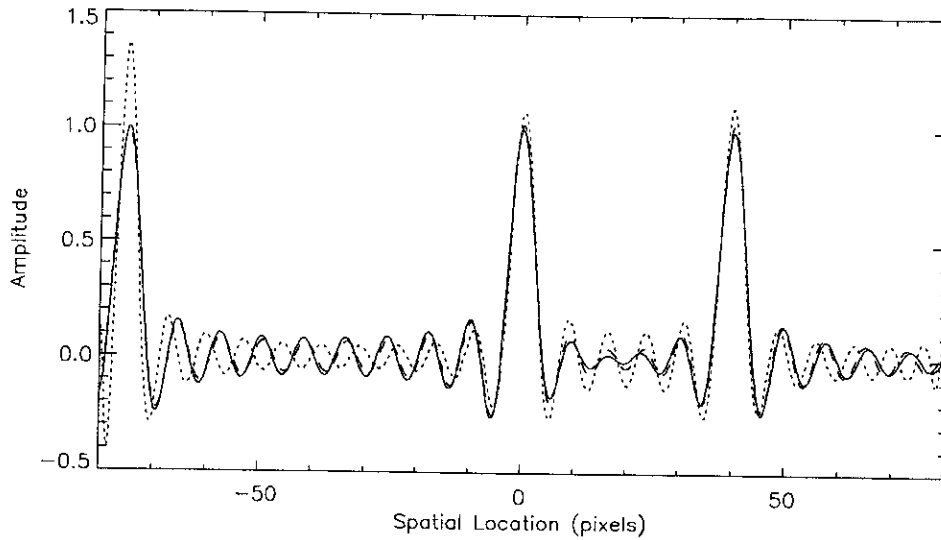

Fig. 13. Plots of reconstructions of the triple star whose Fourier amplitudes are shown in Fig. 8, where the solid line is a reconstruction using just the Fourier data inside the measurement bandwidth, the dashed line is a reconstruction using both the measured Fourier data and the primary superresolution data, and the dotted line is a reconstruction using the measured Fourier data and both the primary and secondary superresolution data.

6. Discussion and Future Work

The results in the preceding sections provide insight into the properties of superresolved data. It was shown that these properties can be seen more clearly if the superresolved Fourier data region is decomposed into two separate regions. The first region is connected to and outside of the measurement bandwidth and is the region where essentially unbiased, or primary, superresolved data resides. The second region is outside of both of these regions and connected to the primary superresolution region and is the region where biased, or secondary, superresolved data resides. Although the secondary superresolution region is theoretically infinite in size, the data SNR limits it to be finite.

Primary superresolution is additive in nature in that the amount of primary superresolution is independent of the SBP as long as the SBP $\gtrsim$ 10 and depends only on the data SNR. For SBPs on the order of one, the amount of primary superresolution for a given data SNR is less than for larger values of the SBP. The increase in resolution in the image domain is constant across the support region. In addition, for objects with SBPs $\gtrsim$ 10, the amount of primary superresolution achievable with realistic data SNRs is on the order of a few percent or less.

Secondary superresolution produces increases in resolution in the support region that are a function of the location in the support region. Across most of the support region away from its edges the resolution increase due to secondary superresolution is essentially constant and has many of the same properties as primary superresolution (i.e., the amount of resolution increase is independent of the SBP when the SBP $\gtrsim 10$, and for smaller SBPs the amount of resolution increase is less). In addition, secondary superresolution in the interior of the support region adds more resolution to the primary resolution results, but still is on the order of a few percent for data with SBPs $\gtrsim 10$. These properties are due to the portion of the superresolved data in the (small) region of the secondary-superresolution Fourier data adjacent to the primary superresolution region. However, the amount of resolution increase at the edges of the support region due to secondary superresolution is much greater than in the interior of the support region. For example, for data with a SBP of 40, the resolution increase averaged over the edges of the support region is ~30%. These higher resolutions are associated with the long "tails" of the secondary superresolution data in the Fourier domain.

Based upon these properties of primary and secondary superresolution, it is clear that both schools of thought regarding the achievability of meaningful superresolution can be correct depending upon the metric used to quantify the phrase "meaningful superresolution". If the metric regarding the achievability of meaningful superresolution is the obtainment of more than a few percent resolution increase across the entire support region for data with SBP $\gtrsim 10$, meaningful superresolution is not possible. If the metric is obtaining more than a few percent resolution increase at the edges of the support for any SBP, meaningful superresolution is possible.

This latter point is worth emphasizing. It is well known that the human visual system keys off of edges in an image. Therefore, if the object under consideration is an object with well-defined edges on a black background (such as is common in astronomical imaging), meaningful superresolution appears to be quite possible and worth pursuing. On the other hand, if the edges of interest are interior to the object support, little edge enhancement is possible.

The results presented in this paper have been independent of the actual object being imaged and superresolved. It is important to determine how effective primary and secondary superresolution are in improving the resolution of images in terms of looking at actual examples. Our initial results (Fig. 13) indicate that using secondary superresolution in the reconstruction process can increase resolution at the edges of the support region but at the cost of decreased radiometric accuracy. In addition, our results only included noise in the sense of limiting the number of PSWFs that could be included in the reconstruction. We expect that noisy data will require additional constraints in the reconstruction process such as Fourier-domain regularization. For these reasons, we are continuing our research into the image domain effects of secondary superresolution and will report on the results a future publication.

Acknowledgements